NORTH-HOLLAND

On Geometric Properties of the Numerical Range

Mao-Ting Chien,* Lina Yeh, and Yean-Tzuoh Yeh
Department of Mathematics
Soochow University
Taipei, Taiwan 11102

and

Fun-Zin Lin
Lih-Chih School of Technology and Commerce
Kaohsiung, Taiwan 807

Submitted by Michael Neumann

ABSTRACT

We investigate the shape of the numerical range. A criterion for the numerical range of a matrix to be an elliptical disk is given. The result is applied to show that there exist neither 3-by-3 nor 4-by-4 nilpotent matrices whose numerical range is an elliptical (noncircular) disk. Sufficient conditions for n-by-n tridiagonal matrices to have elliptical numerical range are obtained. The boundary of the numerical range near a sharp point is examined. Finally, the numerical range of a reducible matrix is compressed, and its geometric properties are discussed. © 1998 Elsevier Science Inc.

1. INTRODUCTION

Let $A \in M_n(\mathbf{C})$, the n-by-n complex matrices; define the *numerical range* of A to be the set

$$W(A) = \{x^*Ax : x \in \mathbf{C}^n,\ x^*x = 1\}.$$

* The work of this author was supported by the National Science Council of the Republic of China.

655 Avenue of the Americas, New York, NY 10010
0024-3795/98/$19.00
PII S0024-3795(97)00373-X

It is a well-known result due to Toeplitz and Hausdorff that the numerical range is always a convex set. Basic properties and references on the numerical range can be found in [6]. In particular, geometric properties of the numerical range have been studied by several authors, e.g., [1, 4, 5, 8, 11]. Donoghue [5] proved that every sharp point of the numerical range of a matrix is an eigenvalue of the matrix. Conditions for the numerical range to be a circular disk were investigated in [4, 11]. From a geometric standpoint, we give, in Section 2, an ellipticity criterion for the numerical range, and show that neither 3-by-3 nor 4-by-4 nilpotent matrices can have an elliptical (noncircular) numerical range. In Section 3, we find sufficient conditions for an n-by-n tridiagonal matrix so that its numerical range is an elliptical disk. In Section 4, we examine the shape of the numerical range near a sharp point on the boundary. Finally, in Section 5, we compress the numerical range of a reducible matrix and discuss its geometric properties.

2. ELLIPTICITY

Let $A \in M_n(\mathbf{C})$. For each θ, we denote by $H_\theta(A)$ the hermitian part of $e^{i\theta}A$, and also denote by $E_k(H_\theta(A))$ the sum of all k-square principal subdeterminants of the matrix $H_\theta(A)$. Then the characteristic polynomial of $H_\theta(A)$ is equal to

$$t^n + \sum_{k=1}^{n} (-1)^k E_k(H_\theta(A)) t^{n-k}.$$

Marcus and Pesce [11] showed that $W(A)$ is a circular disk centered at the origin if and only if the maximal eigenvalue $\lambda_{\max}(H_\theta(A))$ of $H_\theta(A)$ is independent of θ. This result also gives algebraic criteria for an upper triangular matrix so that its numerical range is a circular disk [4, 11]. Now the question arises: What is $\lambda_{\max}(H_\theta(A))$ if $W(A)$ is an elliptical disk? We obtain the following criterion.

THEOREM 1. *Let $A \in M_n(\mathbf{C})$, and $a, b \in \mathbf{R}$. Then $W(A)$ is an elliptical disk with center at the origin, horizontal major axis of length $2a$, and vertical minor axis of length $2b$ if and only if $\lambda_{\max}(H_\theta(A)) = [a^2 - (a^2 - b^2)\sin^2\theta]^{1/2}$ for $0 \leqslant \theta < 2\pi$.*

Proof. Suppose that $W(A)$ is an elliptical disk with center at the origin, horizontal major axis of length of $2a$, and vertical minor axis length of $2b$.

Then the distance c of the foci from the center is equal to $(a^2 - b^2)^{1/2}$ (here we assume $c \neq 0$). Since $W(e^{i\theta}A)$ is a rotation of $W(A)$ through an angle θ, it follows that the foci of $W(e^{i\theta}A)$ are $(c\cos\theta, c\sin\theta)$ and $(-c\cos\theta, -c\sin\theta)$. Then the equation of the ellipse of the boundary of $W(e^{i\theta}A)$ becomes

$$\left[(x - c\cos\theta)^2 + (y - c\sin\theta)^2\right]^{1/2} + \left[(x + c\cos\theta)^2 + (y + c\sin\theta)^2\right]^{1/2} = 2a,$$

that is,

$$(a^2 - c^2\cos^2\theta)x^2 + (a^2 - c^2\sin^2\theta)y^2 - 2c^2\sin\theta\cos\theta\, xy - a^2b^2 = 0. \tag{1}$$

Next we locate the rightmost point of the ellipse (1) by using the Lagrange method. Define

$$F(x, y, \lambda) = x + \lambda\left[(a^2 - c^2\cos^2\theta)x^2 + (a^2 - c^2\sin^2\theta)y^2 - 2c^2\sin\theta\cos\theta\, xy - a^2b^2\right].$$

We compute that

$$F_x = 1 + \lambda\left[2(a^2 - c^2\cos^2\theta)x - 2c^2y\sin\theta\cos\theta\right] = 0, \tag{2}$$

$$F_y = \lambda\left[2(a^2 - c^2\sin^2\theta)y - 2c^2x\sin\theta\cos\theta\right] = 0, \tag{3}$$

$$F_\lambda = (a^2 - c^2\cos^2\theta)x^2 + (a^2 - c^2\sin^2\theta)y^2 - 2c^2xy\sin\theta\cos\theta - a^2b^2 = 0. \tag{4}$$

From (3), we have $\lambda = 0$ or $2(a^2 - c^2\sin^2\theta)y - 2c^2x\sin\theta\cos\theta = 0$. By (2), $\lambda \neq 0$, hence $2(a^2 - c^2\sin^2\theta)y - 2c^2x\sin\theta\cos\theta = 0$, and thus

$$y = \frac{c^2\sin\theta\cos\theta}{a^2 - c^2\sin^2\theta}x. \tag{5}$$

Substituting (5) into (4), we find that

$$\left(a^2 - c^2 \cos^2 \theta - \frac{c^4 \sin^2 \theta \cos^2 \theta}{a^2 - c^2 \sin \theta}\right) x^2 - a^2 b^2 = 0. \tag{6}$$

From (6) we obtain $x = \pm(a^2 - c^2 \sin^2 \theta)^{1/2}$, and it follows that

$$\lambda_{\max}(H_\theta(A)) = (a^2 - c^2 \sin^2 \theta)^{1/2}. \tag{7}$$

Conversely, if (7) holds, then by [10, Theorem 2.1],

$$l_\theta = \left\{z \in \mathbf{C} : \operatorname{Re} z = \lambda_{\max}(H_\theta(A)) = (a^2 - c^2 \sin^2 \theta)^{1/2}\right\}$$

is the right vertical supporting line of the set $W(e^{i\theta}A)$. We shall compute the envelope of the family of the supporting lines, and the envelope describes the boundary of the numerical range of A. For more detail on computing the envelope, see [3]. Observe that the rectangular coordinate system for the supporting line $e^{-i\theta} l_\theta$ of $W(A)$ is

$$y = \cot \theta \left[x - \cos \theta \,(a^2 - c^2 \sin^2 \theta)^{1/2}\right] - \sin \theta \,(a^2 - c^2 \sin^2 \theta)^{1/2}. \tag{8}$$

Geometrically, Equation (8) shows that $W(A)$ is symmetric with respect to the two axes; consequently, there is no loss of generality in restricting $0 < \theta < \pi/2$. Substituting $s = \sin \theta \in (0, 1)$ in (8), we obtain that

$$sy = (1 - s^2)^{1/2} x - (a^2 - c^2 s^2)^{1/2}.$$

Define the function

$$f(s) = sy - (1 - s^2)^{1/2} x + (a^2 - c^2 s^2)^{1/2}; \tag{9}$$

then $f(s) = 0$ for $s \in (0, 1)$. From (9),

$$f'(s) = y + sx(1 - s^2)^{-1/2} - c^2 s(a^2 - c^2 s^2)^{-1/2}. \tag{10}$$

Solving the equation $f(s) - sf'(s) = 0$, we obtain that

$$x = a^2(1 - s^2)(a^2 - c^2 s^2)^{-1/2}. \tag{11}$$

Substituting (11) into the equation $f'(s) = 0$ of (10), we find that

$$y = -s(a^2 - c^2)(a^2 - c^2s^2)^{-1/2}. \tag{12}$$

From (11) and (12), we readily get

$$\frac{x^2}{a^2} + \frac{y^2}{a^2 - c^2} = 1,$$

and this completes the proof. ■

If $A \in M_n(\mathbf{C})$ is a nilpotent upper triangular matrix, Chien and Tam [4] gave a necessary and sufficient condition so that $W(A)$ is a circular disk for $n = 3$ or 4. As a consequence of Theorem 1, however, we show in the following that this class of matrices cannot have an elliptical (noncircular) numerical range.

COROLLARY 1. *If $A \in M_n(\mathbf{C})$ with $n = 3$ or $n = 4$ is a nilpotent upper triangular matrix, then $W(A)$ is not an elliptical disk centered at the origin unless it degenerates to a circular disk or to a single point.*

Proof. Suppose $W(A)$ is an elliptical (noncircular) disk with center at the origin, horizontal major axis of length $2a$, and vertical minor axis of length $2b$. For $n = 3$, the characteristic polynomial of $H_\theta(A)$ is

$$P_{3,\theta}(t) = t^3 - E_1(H_\theta(A))t^2 + E_2(H_\theta(A))t - E_3(H_\theta(A)),$$

where

$$E_1(H_\theta(A)) = 0, \qquad E_2(H_\theta(A)) = s_1, \qquad E_2(H_\theta(A)) = s_2e^{i\theta} + \bar{s}_2e^{-i\theta},$$

and s_1, s_2 are complex constants. By (7), $l(\theta) = (a^2 - c^2\sin^2\theta)^{1/2}$ is the maximal eigenvalue of $H_\theta(A)$ and is satisfied by $P_{3,\theta}(l(\theta)) = 0$ for all θ. Observe that $\sin^2\theta = [a^2 - l(\theta)^2]/c^2$ and $\cos^2\theta = [c^2 - a^2 + l(\theta)^2]/c^2$.

Substituting these observations into $P_{3,\theta}(l(\theta)) = 0$, we compute that

$$
\begin{aligned}
&l(\theta)^3 + s_1 l(\theta) - s_2 e^{i\theta} - \bar{s}_2 e^{-i\theta} \\
&\quad = l(\theta)^3 + s_1 l(\theta) - (s_2 + \bar{s}_2)\cos\theta - (is_2 - i\bar{s}_2)\sin\theta \\
&\quad = l(\theta)^3 + s_1 l(\theta) - (s_2 + \bar{s}_2)\frac{\left[c^2 - a^2 + l(\theta)^2\right]^{1/2}}{c} \\
&\qquad -(is_2 - i\bar{s}_2)\frac{\left[a^2 - l(\theta)^2\right]^{1/2}}{c} = 0.
\end{aligned} \tag{13}
$$

We rewrite the last equality of (13) in the following form:

$$
l(\theta)^3 + p_1 l(\theta) = p_2\left[p_3 l(\theta)^2 + p_4\right]^{1/2} + p_5\left[p_6 l(\theta)^2 + p_7\right]^{1/2}, \tag{14}
$$

where p_i are complex constants. Squaring both sides of Equation (14), we have

$$
\begin{aligned}
&l(\theta)^6 + 2p_1 l(\theta)^4 + p_1^2 l(\theta)^2 \\
&\quad = p_2^2\left[p_3 l(\theta)^2 + p_4\right) + p_5^2\left(p_6 l(\theta)^2 + p_7\right) \\
&\qquad + 2p_2 p_5\left(p_3 l(\theta)^2 + p_4\right)\left(p_6 l(\theta)^2 + p_7\right)\Big]^{1/2}.
\end{aligned} \tag{15}
$$

Moving the first two terms in the right-hand side of (15) to the left-hand side, and then squaring both sides of the equation we have,

$$
\begin{aligned}
&l(\theta)^{12} + u_1 l(\theta)^{10} + u_2 l(\theta)^8 + u_3 l(\theta)^7 + u_4 l(\theta)^6 + u_5 l(\theta)^5 + u_6 l(\theta)^4 \\
&\qquad + u_7 l(\theta)^3 + u_8 l(\theta)^2 + u_9 l(\theta) + u_{10} = 0,
\end{aligned} \tag{16}
$$

where u_i are complex constants. But now Equation (16) has infinite roots $l(\theta)$, which causes a contradiction.

Similarly, for $n = 4$, we compute the characteristic polynomial of $H_\theta(A)$:

$$P_{4,\theta}(t) = t^4 - E_1(H_\theta(A))t^3 + E_2(H_\theta(A))t^2 - E_3(H_\theta(A))t + E_4(H_\theta(A)),$$

where

$$E_1(H_\theta(A)) = 0, \qquad E_2(H_\theta(A)) = s_1, \qquad E_3(H_\theta(A)) = s_2 e^{i\theta} + \bar{s}_2 e^{-i\theta},$$
$$E_4(H_\theta(A)) = s_3 e^{2i\theta} + \bar{s}_3 e^{-2i\theta} + s_4,$$

and s_1, s_2, s_3, s_4 are complex constants. Then for any θ

$$P_{4,\theta}(l(\theta)) = l(\theta)^4 + s_1 l(\theta)^2 - \left(s_2 e^{i\theta} + \bar{s}_2 e^{-i\theta}\right) l(\theta) + s_3 e^{2i\theta} + \bar{s}_3 e^{-2i\theta} + s_4 = 0. \tag{17}$$

Rearranging Equation (17), we obtain that

$$l(\theta)^4 + p_1 l(\theta)^2 + p_2 \cos\theta\, l(\theta) + p_3 \sin\theta\, l(\theta) + p_4 \cos 2\theta + p_5 \sin 2\theta + p_6 = 0, \tag{18}$$

for some constants p_i. Expressing (18) in terms of a, c, and $l(\theta)$, we have

$$l(\theta)^4 + p_1 l(\theta)^2 + p_2 l(\theta)\left[c^2 - a^2 + l(\theta)^2\right]^{1/2} + p_3 l(\theta)\left[a^2 - l(\theta)^2\right]^{1/2} + p_4\left(c^2 - a^2 + l(\theta)^2 - \frac{c^2}{2}\right) + p_5\left[a^2 - l(\theta)^2\right]\left[c^2 - a^2 + l(\theta)^2\right]^{1/2} + p_6 = 0. \tag{19}$$

Then we treat Equation (19) in a similar way by collecting terms and squaring; we find that

$$l(\theta)^{32} + [\text{polynomial in } l(\theta) \text{ of degree less than } 32] = 0,$$

which has infinite roots, once again a contradiction. ■

The result of Corollary 1 indicates that we might expect an extension to general n. But when $n = 5$, it seems fairly complicated to apply the argument in the proof of Corollary 1 to the equation $P_{5,\theta}(l(\theta)) = 0$. Extension of Corollary 1 to general n seems worthy of further study.

3. TRIDIAGONAL MATRICES

A matrix $A = [a_{i,j}] \in M_n(\mathbf{C})$ is *tridiagonal* if $a_{i,j} = 0$ whenever $|i - j| > 1$.Chien [2] showed that the numerical range of a tridiagonal matrix with 0 main diagonal is symmetric with respect to the origin. Eiermann [6, Corollary 4] proved that $W(A)$ is an elliptical disk if $A = [a_{i,j}] \in M_n(\mathbf{C})$ with $a_{i,i+1} = a$ and $a_{i+1,i} = b$ for $i = 1, 2, \ldots, n-1$ and 0 elsewhere for some complex numbers a and b. Indeed, the elliptical disk is explicitly described as

$$\left\{ z \in \mathbf{C} : z = \cos\left(\frac{\pi}{n+1}\right)\left(ae^{-i\theta} + be^{i\theta}\right), 0 \leqslant \theta < 2\pi \right\}.$$

For a matrix $A \in M_n(\mathbf{C})$, given a set $J \subset \{1, 2, 3, \ldots, n\}$, we denote by $A(J')$ the principal submatrix of A formed by deleting rows and columns indicated by J. Let $A = [a_{i,j}] \in M_n(\mathbf{C})$ be a tridiagonal matrix with 0 main diagonal. For every θ, the characteristic polynomial $P_{n,\theta}(t)$ of $H_\theta(A)$ satisfies the recurrence (see, for instance, [2, p. 209])

$$P_{n,\theta}(t) = tP_{n-1,\theta}(t) - \tfrac{1}{4}|a_{1,2}e^{i\theta} + \bar{a}_{2,1}e^{-i\theta}|^2 P_{n-2,\theta}(t), \qquad (20)$$

where $P_{n-1,\theta}(t) = \det([tI - H_\theta(A)](\{1\}'))$ and $P_{n-2,\theta}(t) = \det([tI - H_\theta(A)]\{1,2\}')$.

In the following, we generalize Eiermann's result [6, Corollary 4].

THEOREM 2. *Let $A = [a_{i,j}] \in M_n(\mathbf{C})$ be a tridiagonal matrix with 0 main diagonal. If there exist complex numbers a and b with $|a_{i,i+1}| + |a_{i+1,i}| = |a| + |b|$ and $a_{i,i+1}a_{i+1,i} = ab$ for $i = 1, 2, \ldots, n-1$, then $W(A)$ is an elliptical disk (possibly degenerate) centered at the origin.*

Proof. Let $P_{n,\theta}(t) = \det[tI - H_\theta(A)]$. Then the recurrence (20) holds for $P_{n,\theta}(t)$. In addition

$$\begin{aligned} |a_{1,2}e^{i\theta} + \bar{a}_{2,1}e^{-i\theta}|^2 &= |a_{1,2}|^2 + |a_{2,1}|^2 + a_{1,2}a_{2,1}e^{2i\theta} + \bar{a}_{1,2}\bar{a}_{2,1}e^{-2i\theta} \\ &= (|a| + |b|)^2 - 2|ab| + abe^{2i\theta} + \overline{ab}\, e^{-2i\theta}. \end{aligned}$$

Thus, the characteristic polynomial of $H_\theta(A)$ becomes

$$P_{n,\theta}(t) = tP_{n-1,\theta}(t)$$
$$-\tfrac{1}{4}\Big[(|a| + |b|)^2 - 2|ab| + abe^{2i\theta} + \overline{ab}\, e^{-2i\theta}\Big]P_{n-2,\theta}(t). \quad (21)$$

Since the polynomial (21) is exactly the same as the characteristic polynomial of $H_\theta(B)$, where $B = [b_{i,j}] \in M_n(\mathbf{C})$ is a tridiagonal matrix such that $b_{i,i+1} = a$ and $b_{i+1,i} = b$ for all i, and 0 elsewhere [6, Corollary 4], it follows that the supporting lines $\{z : \operatorname{Re} z = \max w,\ w \in W(e^{i\theta}A)\}$ and $\{z : \operatorname{Re} z = \max w,\ w \in W(e^{i\theta}B)\}$ are identically equal to the line $\{z : \operatorname{Re} z = \text{maximal zero of (21)}\}$. Consequently, by [11, p. 146], $W(A) = W(B)$ and is an elliptical disk centered at the origin. ∎

EXAMPLE 1. Consider matrices

$$A = \begin{bmatrix} 0 & a & 0 & 0 & 0 \\ b & 0 & b & 0 & 0 \\ 0 & a & 0 & -b & 0 \\ 0 & 0 & -a & 0 & -a \\ 0 & 0 & 0 & -b & 0 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 0 & a & 0 & 0 & 0 \\ b & 0 & a & 0 & 0 \\ 0 & b & 0 & a & 0 \\ 0 & 0 & b & 0 & a \\ 0 & 0 & 0 & b & 0 \end{bmatrix}.$$

According to Theorem 2, both $W(A)$ and $W(B)$ are elliptical disks with center at the origin. Moreover, it is easy to see that $W(A) = W(B)$.

Theorem 1 allows us to generalize Eiermann's result in the following way.

THEOREM 3. *Let $A = [a_{i,j}] \in M_n(\mathbf{R})$ be a real tridiagonal matrix with 0 main diagonal. If $a_{i,i+1}a_{i+1,i} > 0$ and there exists $\gamma \in \mathbf{R}$ with $(a_{i,i+1} + a_{i+1,i})^2/(4a_{i,i+1}a_{i+1,i}) = \gamma$ for $i = 1, 2, \dots, n-1$, then $W(A)$ is an elliptical disk with center at the origin, horizontal major axis of length $2a$, and focus at c satisfying $\gamma c^2 = a^2$.*

Proof. Let $P_{n,\theta}(t)$ be the characteristic polynomial of $H_\theta(A)$. From (20), we have the recurrence

$$P_{n,\theta}(t) = tP_{n-1,\theta}(t) - \tfrac{1}{4}b_{1,2}^2 P_{n-2,\theta}(t),$$

where

$$\begin{aligned} b_{j,j+1}^2 &= |a_{j,j+1}e^{i\theta} + \bar{a}_{j+1,j}e^{-i\theta}|^2 \\ &= 2a_{j,j+1}a_{j+1,j}\cos 2\theta + a_{j,j+1}^2 + a_{j+1,j}^2 \\ &= (a_{j,j+1} + a_{j+1,j})^2 - 4a_{j,j+1}a_{j+1,j}\sin^2\theta \\ &= 4a_{j,j+1}a_{j+1,j}(\gamma - \sin^2\theta). \end{aligned}$$

When $n = 2m$ is even, the characteristic polynomial of $H_\theta(A)$ becomes

$$\begin{aligned} P_{2m,\theta}(t) &= t^{2m} + (-1)^1\left(\tfrac{1}{2}\right)^2 t^{2m-2}\sum_{r_1} b_{r_1,r_1+1}^2 \\ &\quad + (-1)^2\left(\tfrac{1}{2}\right)^4 t^{2m-4}\sum_{\substack{r_1<r_2,\\ r_2-r_1\geqslant 2}} b_{r_1,r_1+1}^2 b_{r_2,r_2+1}^2 + \cdots \\ &\quad + (-1)^m\left(\tfrac{1}{2}\right)^{2m}\sum_{\substack{r_1<r_2<\cdots<r_m,\\ r_{j+1}-r_j\geqslant 2}} b_{r_1,r_1+1}^2 b_{r_2,r_2+1}^2\cdots b_{r_m,r_m+1}^2, \end{aligned}$$

where $r_1, r_2, \dots, r_m \in \{1, 2, \dots, 2m-1\}$. Then we compute that

$$\begin{aligned} P_{2m,\theta}(t) &= t^{2m} + (-1)^1(\gamma - \sin^2\theta)^1 t^{2m-2}\sum_{r_1} a_{r_1,r_1+1}a_{r_1+1,r_1} \\ &\quad + (-1)^2(\gamma - \sin^2\theta)^2 t^{2m-4} \\ &\quad \times \sum_{\substack{r_1<r_2,\\ r_2-r_1\geqslant 2}} a_{r_1,r_1+1}a_{r_1+1,r_1}a_{r_2,r_2+1}a_{r_1+1,r_2} + \cdots \\ &\quad + (-1)^m(\gamma - \sin^2\theta)^m \\ &\qquad \times \sum_{\substack{r_1<r_2<\cdots<r_m,\\ r_{j+1}-r_j\geqslant 2}} a_{r_1,r_1+1}a_{r_1+1,r_1}\cdots a_{r_m,r_m+1}a_{r_m+1,r_m} \\ &= t^{2m} + z_{2m-2}(\gamma - \sin^2\theta)^1 t^{2m-2} \\ &\quad + z_{2m-4}(\gamma - \sin^2\theta)^2 t^{2m-4} + \cdots \\ &\quad + z_2(\gamma - \sin^2\theta)^{m-1}t^2 + z_0(\gamma - \sin^2\theta)^m, \end{aligned} \tag{22}$$

where z_i are real constants. Assume the $2m$ real roots of $P_{2m,\theta}(t) = 0$ are $t_1 \geqslant t_2 \geqslant \cdots \geqslant t_{2m}$. Consider the equation with coefficients from (22),

$$x^{2m} + z_{2m-2}x^{2m-2} + z_{2m-4}x^{2m-4} + \cdots + z_2 x^2 + z_0 = 0. \tag{23}$$

Then the $2m$ real roots $x_1 \geqslant x_2 \geqslant \cdots \geqslant x_{2m}$ of (23) satisfy the relation $t_i = x_i(\gamma - \sin^2\theta)^{1/2}$, $i = 1, 2, \dots, 2m$. Now, we choose $c = x_1$, set $a^2 = c^2\gamma$, and set $\mu = (a^2 - c^2\sin^2\theta)^{1/2}$. It is easy to verify that $\gamma - \sin^2\theta = \mu^2/c^2$, and thus (22) becomes

$$P_{2m,\theta}(t) = t^{2m} + z_{2m-2}(\mu/c)^2 t^{2m-2} + z_{2m-4}(\mu/c)^4 t^{2m-4} + \cdots$$
$$+ z_2(\mu/c)^{2m-2}t^2 + z_0(\mu/c)^{2m}.$$

It follows that

$$P_{2m,\theta}(\mu) = (\mu/c)^{2m}\big(c^{2m} + z_{2m-2}c^{2m-2}$$
$$+ z_{2m-4}c^{2m-4} + \cdots + z_2c^2 + z_0\big)\mu_m^2 = 0.$$

Hence μ is an eigenvalue of $H_\theta(A)$. Moreover, $t_i = x_i\mu/c = x_i\mu/x_1 \leqslant \mu$; therefore, μ is the maximal eigenvalue of $H_\theta(A)$, and the result now follows from Theorem 1. The case when $n = 2m + 1$ can be treated in a similar way. ■

Similarly, we have

THEOREM 4. *Let $A = [a_{i,j}] \in M_n(\mathbf{R})$ be a real tridiagonal matrix with 0 main diagonal. If $a_{i,i+1}a_{i+1,i} < 0$ and there exists $\gamma \in \mathbf{R}$ with $(a_{i,i+1} + a_{i+1,i})^2/(-4a_{i,i+1}a_{i+1,i}) = \gamma$ for $i = 1, 2, \dots, n-1$, then $W(A)$ is an elliptical disk with center at the origin, vertical major axis of length $2a$, and focus at c satisfying $\gamma c^2 = a^2 - c^2$.*

Proof. In this case,

$$b_{j,j+1}^2 = -4a_{j,j+1}a_{j+1,j}(\gamma + \sin^2\theta).$$

We choose $c = x_1$ and $a^2 = \gamma c^2 + c^2$ and let $\mu = [a^2 - c^2\sin^2(\theta + \pi/2)]^{1/2} = (a^2 - c^2\cos^2\theta)^{1/2}$. The rest of the proof follows from an argument used in the proof of Theorem 3. ■

We remark that the focus of the elliptical disk in Theorem 3 can be found by computing the maximal root of Equation (23). For instance, it is not difficult to determine that $n = 3$,

$$c^2 = \frac{1}{2}\Big\{(a_{1,2}a_{2,1} + a_{2,3}a_{3,2} + a_{3,4}a_{4,3}) + \Big[(a_{1,2}a_{2,1} + a_{2,3}a_{3,2} + a_{3,4}a_{4,3})^2 - 4a_{1,2}a_{2,1}a_{3,4}a_{4,3}\Big]^{1/2}\Big\},$$

and for $n = 5$,

$$c^2 = \frac{1}{2}\Big\{(a_{1,2}a_{2,1} + a_{2,3}a_{3,2} + a_{3,4}a_{4,3} + a_{4,5}a_{5,4}) + \Big[(a_{1,2}a_{2,1} + a_{2,3}a_{3,2} + a_{3,4}a_{4,3} + a_{4,5}a_{5,4})^2 - 4(a_{1,2}a_{2,1}a_{3,4}a_{4,3} + a_{1,2}a_{2,1}a_{4,5}a_{5,4} + a_{2,3}a_{3,2}a_{4,5}a_{5,4})\Big]^{1/2}\Big\}.$$

4. SPECIAL GEOMETRIC PROPERTIES

It is clear that for any two ellipses there exists a 4-by-4 matrix whose numerical range is the convex hull of the two ellipses. However, no 3-by-3 matrix has this property. An example is provided in the following.

EXAMPLE 2. Consider the matrix $A = A_1 \oplus A_2$, where

$$A_1 = \begin{bmatrix} 0 & 2 \\ 0 & 0 \end{bmatrix}, \qquad A_2 = \begin{bmatrix} 3 & 2 \\ 0 & 3 \end{bmatrix}.$$

Suppose there exists

$$B = \begin{bmatrix} x & a & b \\ 0 & y & c \\ 0 & 0 & z \end{bmatrix} \in M_3(\mathbf{C})$$

such that $W(B) = W(A)$. By straightforward computations, we obtain the characteristic polynomial of $H_\theta(B)$:

$$P_{3,\theta}(t) = t^3 - E_1(H_\theta(B))t^2 + E_2(H_\theta(B))t - E_3(H_\theta(B)),$$

where

$$E_1(H_\theta(B)) = \tfrac{1}{2}\Big[(x+y+z)e^{i\theta} + \overline{(x+y+z)}\,e^{-i\theta}\Big],$$

$$E_2(H_\theta(B)) = \tfrac{1}{4}\Big[(xy+yz+xz)e^{2i\theta} + \overline{(xy+yz+xz)}\,e^{-2i\theta}$$

$$+2\,\mathrm{Re}\big(x\bar{y}+y\bar{z}+z\bar{x}\big) - \big(|a|^2+|b|^2+|c|^2\big)\Big],$$

$$E_3(H_\theta(B)) = \tfrac{1}{8}\Big[xyze^{3i\theta} + \overline{xyz}\,e^{-3i\theta} + \big(xy\bar{z}+x\bar{y}z+\bar{x}yz+a\bar{b}c\big)e^{i\theta}$$

$$+\overline{\big(xy\bar{z}+x\bar{y}z+\bar{x}yz+a\bar{b}c\big)}\,e^{-i\theta}$$

$$-|a|^2\big(ze^{i\theta}+\bar{z}e^{-i\theta}\big) - |b|^2\big(ye^{i\theta}+\bar{y}e^{-i\theta}\big)$$

$$-|c|^2\big(xe^{i\theta}+\bar{x}e^{-i\theta}\big)\Big].$$

Observe that $W(A_1)$ is a unit circular disk, and that 1 is the maximal eigenvalue of $H_\theta(B)$ whenever $\theta \in [\pi/2, 3\pi/2]$; it follows that $P_{3,\theta}(1) = 0$ for $\theta \in [\pi/2, 3\pi/2]$. Then

$$P_{3,\theta}(1) = 1 - E_1(H_\theta(B)) + E_2(H_\theta(B)) - E_3(H_\theta(B))$$

$$= s_0 + s_1e^{i\theta} + \bar{s}_1e^{-i\theta} + s_2e^{2i\theta} + \bar{s}_2e^{-2i\theta} + s_3e^{3i\theta} + \bar{s}_3e^{-3i\theta} = 0 \tag{24}$$

for all $\theta \in [\pi/2, 3\pi/2]$, where s_0, s_1, s_2, s_3 are complex constants. By (24), $s_0 = s_1 = s_2 = s_3 = 0$. On the other hand, we compute that

$$s_2 = \frac{xy+yz+xz}{4} = 0, \tag{25}$$

$$s_3 = xyz/8 = 0. \tag{26}$$

Next, for $\theta = [-\pi/2, \pi/2]$, the maximal eigenvalue of $H_\theta(B - 3I)$ is equal to 1 because $W(A_2 - 3I)$ is a unit circular disk. Therefore, $Q_{3,\theta}(1) = 0$ for all $\theta \in [-\pi/2, \pi/2]$, where $Q_{3,\theta}(t)$ is the characteristic polynomial of

$H_\theta(B - 3I)$. By a similar argument to the one given above, we compute the coefficients of $e^{2i\theta}$ and $e^{3i\theta}$ in the equation $Q_{3,\theta}(1) = 0$, which are equal to zero, and we get

$$\tfrac{1}{4}[(x-3)(y-3)+(y-3)(z-3)+(x-3)(z-3)] = 0, \quad (27)$$

$$\tfrac{1}{8}(x-3)(y-3)(z-3) = 0. \quad (28)$$

In view of (25)–(27), we find that $x + y + z = \frac{9}{2}$. On the other hand, by (25), (26), and (28), $x + y + z = 3$, a contradiction. This gives an example to show that there doesn't exist a matrix $B \in M_3(\mathbf{C})$ such that $W(B) = W(A)$.

Let $A \in M_n(\mathbf{C})$. A point $z \in \partial W(A)$ at which $\partial W(A)$ is not differentiable is called a *sharp point* of $W(A)$. Donoghue [5, Theorem 7] proved that every sharp point of $W(A)$ is an eigenvalue of A. We examine the boundary near a sharp point.

THEOREM 5. *Let $A \in M_n(\mathbf{C})$ and $z \in \mathbf{C}$. Then z is a sharp point of $W(A)$ if and only if A is unitarily equivalent to $zI_m \oplus A_1$ with $z \notin W(A_1)$. In this case, z is the vertex of the intersection of two line segments on $\partial W(A)$.*

Proof. It is clear that z is a sharp point if A is unitarily equivalent to $zI_m \oplus A_1$ with $z \notin W(A_1)$. If z is a sharp point of $W(A)$, by [5, Theorem 7] it is an eigenvalue of A. Then there exists a unitary matrix U such that

$$U^*AU = \begin{bmatrix} z & b_{1,2} & \cdots & b_{1,n} \\ & \lambda_2 & & \\ & 0 & \ddots & * \\ & & & \lambda_n \end{bmatrix}.$$

Let

$$B = \begin{bmatrix} z & b_{1,i} \\ 0 & \lambda_i \end{bmatrix}$$

be a 2-by-2 principal submatrix of U^*AU determined by the indices 1 and i, $2 \leqslant i \leqslant n$. If $b_{1,i} \neq 0$ for some i, then z is a focus of the nondegenerate elliptical disk $W(B) \subset W(A)$, which contradicts $z \in \partial W(A)$, and thus $b_{1,i}$

$= 0$ for all i. Assume the multiplicity of z is m. Then A is unitarily equivalent to $zI_m \oplus A_1$ for some $A_1 \in M_{n-m}(\mathbf{C})$ and $z \notin \sigma(A_1)$. Suppose that $z \in W(A_1)$; we have

$$W(A) = \mathrm{Co}(\{z\} \cup W(A_1)) = W(A_1). \tag{29}$$

Thus by (29) z is a sharp point of $W(A_1)$, but then z is an eigenvalue of A_1, a contradiction. ■

It has been an open problem to characterize compact convex sets which are numerical ranges of matrices (see, e.g., [9]). Theorem 5 provides compact convex sets which are not the numerical ranges of matrices, for example, the upper half unit circular disk.

Next, we determine unitary equivalence of matrices which have the same numerical range and have sufficient sharp points.

COROLLARY 2. *Let $A, B \in M_n(\mathbf{C})$ with $W(A) = W(B)$. Then A and B are unitarily equivalent if one of the following conditions is satisfied:*

(i) *$W(A)$ has n sharp points.*
(ii) *$W(A)$ has $n-1$ sharp points and* $\operatorname{tr} A = \operatorname{tr} B$.
(iii) *$W(A)$ has $n-2$ sharp points and is not a polygon.*

Proof. (i): If $W(A) = W(B)$ and has n sharp points $z_1, \dots, z_n$, then by Theorem 5 both A and B are unitarily equivalent to $\mathrm{diag}(z_1, \dots, z_n)$.

(ii): Suppose $W(A) = W(B)$ and has sharp points $z_1, \dots, z_{n-1}$. By the proof of Theorem 5, A and B are unitarily equivalent to $D \oplus [\beta]$ and $D \oplus [\gamma]$ respectively, where $D = \mathrm{diag}(z_1, \dots, z_{n-1})$. Consequently, A and B are unitarily equivalent if and only if $\beta = \gamma$, or equivalently, $\operatorname{tr} A = \operatorname{tr} B$.

(iii): If $W(A) = W(B)$ is not a polygon and has $n-2$ sharp points $z_1, \dots, z_{n-2}$, then A and B are unitarily equivalent to $D \oplus A_1$ and $D \oplus B_1$ respectively, where $D = \mathrm{diag}(z_2, \dots, z_{n-2})$, and $A_1, B_1 \in M_2(\mathbf{C})$. Observe now that

$$W(A) = \mathrm{Co}(W(D) \cup W(A_1)) = \mathrm{Co}(W(D) \cup W(B_1)) = W(B).$$

Since the two ellipses $W(A_1) \not\subset W(D)$ and $W(B_1) \not\subset W(D)$, it follows that $W(A_1) = W(B_1)$. Suppose λ and μ are the foci of the common ellipse. Then A_1 and B_1 are unitarily equivalent to

$$\begin{bmatrix} \lambda & \alpha \\ 0 & \mu \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} \lambda & \beta \\ 0 & \mu \end{bmatrix}$$

respectively. Computing the length of the minor axis of the ellipse, we have

$$\left(\|A_1\| - |\lambda|^2 - |\mu|^2\right)^{1/2} = \left(\|B_1\| - |\lambda|^2 - |\mu|^2\right)^{1/2}.$$

This implies that $|\alpha| = |\beta|$. Thus A_1 and B_1 are unitarily equivalent, and the theorem follows. ∎

5. COMPRESSIONS OF NUMERICAL RANGE

Marcus and Pesce [11] compressed $W(A)$ as the union of the numerical ranges of 2-by-2 compressions. They showed $W(A) = \bigcup_{x,y} W(A_{xy})$, where x, y run over all pairs of real orthonormal vectors, and the 2-by-2 compression matrices

$$A_{xy} = \begin{bmatrix} (Ax, x) & (Ay, x) \\ (Ax, y) & (Ay, y) \end{bmatrix}.$$

This result gives an algorithm for plotting the numerical range of A by exhibiting elliptical disks $W(A_{xy})$ [11, Section III]. In this section, we compress the numerical range of a reducible matrix, and give its geometric properties.

The following result compresses the numerical range into the union of 2-by-2 compressions over orthogonally complemental subspaces.

THEOREM 6 [2, Theorem 1]. *Let $A \in M_n(\mathbf{C})$, and S be a subspace of $\mathbf{C}^n$. Then*

$$W(A) = \bigcup_{x,y} W(A_{xy}),$$

where x and y vary over all unit vectors in S and $S^{\perp}$ respectively.

Theorem 6 leads to the following corollaries.

COROLLARY 3. *Let $A \in M_n(\mathbf{C})$. If $A^2 = 0$, then $W(A)$ is a circular disk centered at the origin.*

Proof. Choose S = kernel A. It is easy to see that the compression matrix A_{xy} in Theorem 6 becomes

$$\begin{bmatrix} 0 & (Ay, x) \\ 0 & 0 \end{bmatrix},$$

whose numerical range is a circular disc centered at the origin with radius $|(Ay, x)|/2$. ■

A matrix $A \in M_n(\mathbf{C})$ is said to be *reducible* if either $n = 1$ and $A = 0$ or $n \geqslant 2$ and there exists a permutation matrix $P \in M_n$ such that

$$P^T A P = \begin{bmatrix} B & C \\ 0 & D \end{bmatrix}, \tag{30}$$

where $B \in M_k(\mathbf{C})$ and $D \in M_{n-k}(\mathbf{C})$, $1 \leqslant k \leqslant n - 1$. Since the numerical range of a matrix is unchanged under a unitary similarity, we may assume that reducible matrix is in the 2-by-2 block-triangular form (30).

COROLLARY 4. *Let*

$$A = \begin{bmatrix} B & C \\ 0 & D \end{bmatrix} \in M_n(\mathbf{C}) \qquad \textit{with} \quad B \in M_k(\mathbf{C}) \ \textit{and} \ D \in M_{n-k}(\mathbf{C}).$$

Then

$$W(A) = \bigcup_{x,\, y} W\left(\begin{bmatrix} (Bx, x) & (Cy, x) \\ 0 & (Dy, y) \end{bmatrix} \right),$$

where x, y *runs over all unit vectors in* $\mathbf{C}^k$ *and* $\mathbf{C}^{n-k}$ *respectively.*

Proof. Let S be the subspace of $\mathbf{C}^n$ generated by $e_1, \dots, e_k$, where $\{e_1, e_2, \dots, e_n\}$ is the standard basis of $\mathbf{C}^n$. For unit vectors $x \in S$ and $y \in S^{\perp}$, denote $x = [\tilde{x} \ \ 0]^T$ and $y = [0 \ \ \tilde{y}]^T$, where $\tilde{x} \in \mathbf{C}^k$ and $\tilde{y} = \mathbf{C}^{n-k}$. Direct calculations show that

$$\begin{bmatrix} (Ax, x) & (Ay, x) \\ (Ax, y) & (Ay, y) \end{bmatrix} = \begin{bmatrix} (B\tilde{x}, \tilde{x}) & (C\tilde{y}, \tilde{x}) \\ 0 & (D\tilde{y}, \tilde{y}) \end{bmatrix}.$$

By Theorem 6,

$$W(A) \subset \bigcup_{x,y} W\left(\begin{bmatrix} (Bx,x) & (Cy,x) \\ 0 & (Dy,y) \end{bmatrix}\right),$$

where $x \in \mathbf{C}^k$ and $y \in \mathbf{C}^{n-k}$ are unit vectors. Conversely, let $x \in \mathbf{C}^k$, $y \in \mathbf{C}^{n-k}$ be unit vectors, and denote $\hat{x} = [x \;\; 0]^T \in \mathbf{C}^n$ and $\hat{y} = [0 \;\; y]^T \in \mathbf{C}^n$. Then the unit vector $\hat{x} \in S$ (the subspace generated by $e_1, \dots, e_k$ in $\mathbf{C}^n$) and similarly $\hat{y} \in S^{\perp}$ is unital, and

$$\begin{bmatrix} (Bx,x) & (Cy,x) \\ 0 & (Dy,y) \end{bmatrix} = \begin{bmatrix} (A\hat{x},\hat{x}) & (A\hat{y},\hat{x}) \\ (A\hat{x},\hat{y}) & (A\hat{y},\hat{y}) \end{bmatrix}.$$

Again by Theorem 6,

$$W\left(\begin{bmatrix} (Bx,x) & (Cy,x) \\ 0 & (Dy,y) \end{bmatrix}\right) \subset W(A),$$

and the conclusion follows. ■

Corollary 4 enables us to find some geometric properties of the numerical range of a reducible matrix.

COROLLARY 5. *Let*

$$A = \begin{bmatrix} B & C \\ 0 & D \end{bmatrix} \in M_n(\mathbf{C})$$

with $B \in M_k(\mathbf{C})$ *and* $D \in M_{n-k}(\mathbf{C})$, $1 \leqslant k \leqslant n-1$. *Then*:

(i) *If* $\operatorname{rank} C = k$, *then* $\sigma(A)$ *is contained in the interior of* $W(A)$ *and the boundary of* $W(A)$ *is differentiable.*

(ii) *If* $B = \alpha I_k$ *and* $D = \beta I_{n-k}$, *then* $W(A)$ *is an elliptical disk with foci* α *and* β, *semimajor axis of length* $\frac{1}{2}(|\alpha|^2 + |\beta|^2 + \|C\|^2 - 2\operatorname{Re}\alpha\bar{\beta})^{1/2}$, *and semiminor axis of length* $\|C\|/2$, *where* $\|C\| = \sup_{|x|=|y|=1}|(Cy,x)|$.

(iii) *If* $n = 2k$ *and* $W(B)$ *is a circular disk centered at the origin with radius* r, *and in addition,* C *is unitary and* $D = 0$, *then* $W(A)$ *is a circular disk centered at the origin with radius* $[r + (r^2+1)^{1/2}]/2$.

(iv) *If* $n = 2k$, $B = D$, *and* $|x^* Cx| = \|C\|$ *whenever* x *is a unit vector* *with* $x^*Bx \in \partial W(B)$, *then* $\operatorname{dist}(\partial W(A), \partial W(B)) \equiv \inf\{|t-u| : t \in \partial W(A), u \in \partial W(B)\} = \|C\|/2$. *Moreover, if* $W(B)$ *is a circular disk centered at the origin with radius* r, *then* $W(A)$ *is a circular disk centered at the origin with radius* $r + \|C\|/2$.

Proof. (i): Let $p \in \sigma(A)$. Then p is an eigenvalue of B or D. Assume $p \in \sigma(B)$. Take a unit vector $x \in \mathbf{C}^k$ such that $p = (Bx, x)$. For this x, choose a unit vector $y \in \mathbf{C}^{n-k}$ such that $(Cy, x) \neq 0$. This can be done because rank $C = k$. It follows that p is one of the foci of the nondegenerate elliptical disk

$$W\left(\begin{bmatrix} (Bx, x) & (Cy, x) \\ 0 & (Dy, y) \end{bmatrix}\right),$$

which by Corollary 4 is contained in $W(A)$. Thus p is an interior point of $W(A)$. A similar argument holds for $p \in \sigma(D)$.

If $W(A)$ has a sharp point z, then by [5], z is an eigenvalue of A. It follows that z is an interior of $W(A)$, a contradiction, and (i) is proved.

If $B = \alpha I_k$ and $D = \beta I_{n-k}$ then for any unit vectors $x \in \mathbf{C}^k$ and $y \in \mathbf{C}^{n-k}$, we have

$$\begin{bmatrix} (Bx, x) & (Cy, x) \\ 0 & (Dy, y) \end{bmatrix} = \begin{bmatrix} \alpha & (Cy, x) \\ 0 & \beta \end{bmatrix}. \tag{31}$$

Then (ii) follows from Corollary 4 and the fact that the numerical range of the matrix on the right-hand side of (31) is an elliptical disk with foci α and β, semimajor axis of length $\frac{1}{2}[|\alpha|^2 + |\beta|^2 + |(Cy, x)|^2 - 2\operatorname{Re}\alpha\bar{\beta}]^{1/2}$, and semiminor axis of length $|(Cy, x)/2|$.

To prove (iii), a typical compression matrix

$$\begin{bmatrix} (Bx, x) & (Cy, x) \\ 0 & (Dy, y) \end{bmatrix}$$

in Corollary 4 reduces to

$$\begin{bmatrix} (Bx, x) & (Cy, x) \\ 0 & 0 \end{bmatrix}.$$

Let x be a unit vector in $\mathbf{C}^k$. Choose a unit vector $z \in \mathbf{C}^{n-k}$ $(= \mathbf{C}^k)$ such that $Cz = x$. This can be achieved, since C is unitary. For any unit vector y we have $|(Cy, x)| \leqslant |Cy|\,|x| \leqslant 1$; it follows that $\sup_{|y|=1}|(Cy, x)| = 1$ for each unit vector x. Then

$$W\left(\begin{bmatrix} (Bx, x) & (Cy, x) \\ 0 & 0 \end{bmatrix}\right) \subset W\left(\begin{bmatrix} (Bx, x) & 1 \\ 0 & 0 \end{bmatrix}\right)$$

for all x and y. Hence

$$W(A) = \bigcup_x W\left(\begin{bmatrix} (Bx, x) & 1 \\ 0 & 0 \end{bmatrix}\right),$$

where x runs over the unit vectors in $\mathbf{C}^k$. Now

$$W\left(\begin{bmatrix} (Bx, x) & 1 \\ 0 & 0 \end{bmatrix}\right) \tag{32}$$

is an elliptical disk with foci 0 and (Bx, x) and semimajor axis of length $\frac{1}{2}[1 + |(Bx, x)|^2]^{1/2}$. Hence for every point (Bx, x) on the circle of radius r, the farthest distance from the ellipse (32) to the original is $[r + (r^2 + 1)^{1/2}]/2$. $W(A)$ is then a circular disk centered at the origin with radius $[r + (r^2 + 1)^{1/2}]/2$. This proves (iii).

To prove (iv), it is clear that $W(B) \subset W(A)$, since B is a principal submatrix of A. For unit vectors x and y,

$$|(Cy, x)|^2 \leqslant \|C\|^2 + 2\|C\||(Bx, x) - (Dy, y)|.$$

Then

$$|(Bx, x) - (Dy, y)|^2 + |(Cy, x)|^2 \leqslant [\|C\| + |(Bx, x) - (Dy, y)|]^2.$$

Computing the length of the semimajor axis of the elliptical disk

$$W\left(\begin{bmatrix} (Bx, x) & (Cy, x) \\ 0 & (Dy, y) \end{bmatrix}\right), \tag{33}$$

we obtain that

$$\begin{aligned}
&\tfrac{1}{2}\Big[|(Bx,x)|^2+|(Dy,y)|^2+|(Cy,x)|^2-2\operatorname{Re}(Bx,x)\overline{(Dy,y)}\Big]^{1/2}\\
&\quad=\tfrac{1}{2}\Big[|(Bx,x)-(Dy,y)|^2+|(Cy,x)|^2\Big]^{1/2}\\
&\quad\leqslant\tfrac{1}{2}\big[\|C\|+|(Bx,x)-(Dy,y)|\big].
\end{aligned}$$

This means that the distance from the endpoints of the semimajor axis of the ellipse (33) to the closest focus is not greater than $\|C\|/2$. The length of the semiminor axis of the ellipse is $|(Cy,x)/2|$, which is not greater than $\|C\|/2$. Moreover, if $x \in \mathbf{C}^k$ is a unit vector such that $(Bx, x) \in \partial W(B)$, pick $y = x$; then

$$W\left(\begin{bmatrix} (Bx,x) & (Cy,x) \\ 0 & (Dy,y) \end{bmatrix}\right)$$

is a circular disk centered at (Bx, x) with radius $|(Cx, x)|/2 = \|C\|/2$. Hence the region $W(A)$ is expanded to a distance $\|C\|/2$ away from the boundary of $W(B)$. Furthermore if $W(B)$ is a circular disk centered at the origin with radius r, then $W(A)$ is a circular disk centered at the origin with radius $r + \|C\|/2$. ∎

6. NOTE

Recently the authors have learned from Hiroshi Nakazato that the result of Corollary 1 can be extended to general n. His idea is based on the work of V. R. Kippenhaha, Über den Wertevorrat einer Matrix, Math. Nachr. 6:193–228 (1951). Suppose that $W(A)$ is a nondegenerated elliptical disk for some nilpotent matrix $A \in M_n(\mathbf{C})$. We may assume without loss of generality that $W(A)$ is an elliptical disk centered at the origin, horizontal semimajor axis of length a and vertical semiminor axis of length b with $a > b$. Then the homogeneous polynomial

$$F(t,x,y) \equiv \det(tI_n + xH + yK)$$

contains a factor $t^2 - a^2x^2 - b^2y^2$, where $A = H + iK$, $H = (A + A^*)/2$ and $K = (A - A^*)/(2i)$. It follows that $F(t, -1, -i) = \det(tI_n - A)$ contains a factor $(t - \sqrt{a^2 - b^2})(t + \sqrt{a^2 - b^2})$, and this contradicts to $\sigma(A) = \{0\}$.

The authors wish to thank the referee for his suggestions which led to improvements in the presentation.

REFERENCES

1 C. S. Ballantine, Numerical range of a matrix, some effective criteria, *Linear Algebra Appl.* 19:117–188 (1978).
2 M. T. Chien, On the numerical range of tridiagonal operators, *Linear Algebra Appl.* 246:203–214 (1996).
3 M. T. Chien, The envelope of the generalized numerical range, *Linear and Multilinear Algebra*, to appear.
4 M. T. Chien and B. S. Tam, Circularity of the numerical range, *Linear Algebra Appl.* 201:113–133 (1994).
5 W. D. Donoghue, On the numerical range of a bounded operator, *Michigan Math. J.* 4:26–263 (1957).
6 M. Eiermann, Fields of values and iterative methods, *Linear Algebra Appl.* 180:167–197 (1993).
7 R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis*, Cambridge U.P., 1991.
8 C. R. Johnson, Numerical determination of the field of values of a general complex matrix, *SIAM J. Numer. Anal.* 15:595–602 (1978).
9 C.-K. Li, C-numerical ranges and C-numerical radii, *Linear and Multilinear Algebra* 37:51–82 (1994).
10 C.-K. Li, C. H. Sung, and N. K. Tsing, c-Convex matrices: Characterizations, inclusion relations and normality, *Linear and Multilinear Algebra* 25:275–287 (1989).
11 M. Marcus and C. Pesce, Computer generated numerical ranges and some resulting theorems, *Linear and Multilinear Algebra* 20:121–157 (1987).

Received 30 August 1996; final manuscript accepted 14 July 1997